1

Introduction: A Brief Review of Methods of Studying Language Production

B. Butterworth *University of Cambridge*

I. Introduction

The properties of human nature that make talk possible have fascinated philosophers since the Enlightenment. For Descartes, animals and machines "could never use speech or other signs as we do when placing our thought on record for others" (1637, p. 116). "Magpies and parrots are able to utter words just like ourselves, and yet they cannot speak as we do, that is, so as to give evidence that they think of what they say" (p. 117); and a machine cannot arrange "its speech in various ways, in order to reply appropriately to everything that may be said in its presence, as even the lowest type of man can do" (p. 116). Thus literally thoughtful talk is incontestable evidence for a fundamental division between human beings and other sublunary creatures, namely, we have a rational soul, they do not.

The role of speech as the medium through which thoughts are conveyed to oneself and to others, and hence a vital component in man's nature as a social, as well as a rational, animal, was recognized by Locke (1700).

> Man, though he have great variety of Thoughts, and such, from which others, as well as himself, might receive Profit and Delight; yet they are all within his own Breast, invisible, and hidden from others, nor can of themselves be made appear. The Comfort, and Advantage of Society, not being to be had without Communication of Thoughts, it was necessary, that Man should find out some external sensible Signs, whereby those invisible *Ideas*, which his thoughts are made of, might be made known to others. For this purpose, nothing was so fit, either for Plenty or Quickness, as those articulate Sounds, which with so much Ease and Variety, he found himself able to make. Thus we may conceive how *Words*, which were by Nature so well adapted to that purpose, come to be made use of by Men, as *the Signs of* their *Ideas*; not by any natural connexion, that there is between particular articulate Sounds and certain *Ideas*, for then there would be but one Language amongst all Men; but by a voluntary Imposition, whereby such a Word is made arbitrarily the Mark of such an Idea. The use then of Words, is to be sensible Marks of *Ideas*. (404–405)

A necessary prolegomenon to his philosophy was thus an analysis of the signification of words, which, for him, meant both an account of how words

come to refer to objects in the world and a proto-psychological treatment of words as the expression of mental entities, "Ideas".

More recently, and more scientifically, the problem of the mental apparatus responsible for speech has attracted research and speculation from some of the most profound students of human nature (Hughlings Jackson, 1958; Freud, 1891, 1924; Wundt, 1900; Pick, 1931; Goldstein, 1948; Lashley, 1951; George Miller, 1960). Even so, modern psychologists of language have, by and large, either ignored the problem or treated the research with scepticism or pessimism. Thus, in their introductory psycholinguistic text, Glucksberg and Danks (1975) devote only two pages to production. Johnson-Laird (1974) has written in a general review of psycholinguistics, "the fundamental problem in psycholinguistics is simple to formulate: what happens when we understand sentences" (p. 135). MacNeilage and Ladefoged (1976), reviewing the "production of speech and language" write: "very little is known about the production of language" (p. 75). And even where a text devotes considerable space to production, we find "practically anything that one can say about speech production must be considered speculative even by the standards current in psycholinguistics" (Fodor *et al.*, 1974, p. 434).

Why should the study of production evoke these expressions of skepticism, pessimism and neglect? One reason seems to be that experimental psychologists like to be able to *manipulate* at least some of the relevant variables and to have some *control* over the range of options available to the subject. Usually, this has meant manipulating the input to the subject and restricting the range of responses the subject is allowed to use. In this way, complex phenomena and behaviours can be subdivided into more manageable components, and systematic input-output relations can be established which will lead to confident inferences about the processes intervening between input and output. Now, what we say typically bears little systematic relationship to environmental input (*pace* Skinner, 1957), and thus it would be extraordinarily optimistic to set up manipulations of the input and expect to find systematic outputs, unless the subject is so limited in what he is allowed to say that generalizations to natural spontaneous speech become almost impossible. So psycholinguistics has concentrated on input-end processing—word-recognition, comprehension and the like—where manipulation of stimuli and the limitation of response choices seems a more plausible and fruitful strategy.

A second reason for avoiding production is that speech occurs naturally mainly in conversations, and these are, in many ways, no less than microcosms of the social order. So, in order to get a grip on what is going on in production, not only must the usual set of psychological variables be taken into account, but so must a new range of social variables. The evident complexity of the phenomena and the difficulty of identifying the responsible variable has, no doubt, deterred many potential investigators (for further discussion of this issue, see Butterworth, 1978).

However, there is a price to pay for control in the study of input-end processes. First, the products of word recognition, comprehension or

whatever, are not directly observable, but must be inferred from behaviour linked to these processes. Since the experimenter has to limit available responses—or else problems of comprehension combine with the problems of production—the subject is required to generate responses not normally and naturally associated with the stimulus, for example, repeating the word(s) presented, pressing a button on hearing (or seeing) a target, etc., and, of course, one thereby encounters the problem of how to generalize from the experimental task to real-life activity. With production, on the other hand, the natural products or expressions of the underlying processes *are* directly observable, namely, the speech uttered. Indeed, one can use material produced with little or no experimental intervention. So the problem of the generality of the findings is reduced at a stroke.

In addition, it turns out in practice difficult to generalize not only to real-life, but even to other experimental paradigms that are apparently very similar. Hence, it is hard to evaluate the theoretical implications of such equivocal studies.

Let us take an example in which the stimuli are fairly naturalistic, easily manipulable and where the response is as simple as can be. There are large numbers of studies where the subject has to monitor a target while listening to a sentence. As soon as the subject detects the target he must press the button. The idea is that the speed of reaction will indicate how much cognitive work comprehension of the sentence demands at the target location: the greater the current cognitive load, the longer it takes the subject to detect and respond to the target. By manipulating or identifying characteristics of the sentence, it should be possible to tease out which hypothetical processes are engaged in comprehension. For example, if embedded constructions, as compared with right-branching constructions, increase latency to a following target, then syntactic analysis of embedding is cognitively more demanding (Foss and Lynch, 1969). Now, targets have been of two sorts: parts of the sentence (a phoneme, syllable or word) or extraneous noises (tones or "clicks"). Response times to sentence-internal targets are reliably slower near the beginnings of sentences than near the ends (Foss, 1969; Shields *et al.*, 1974; Cutler and Foss, 1977), whereas, response times to extraneous targets are reliably faster at the beginning than at the end (Abrams and Bever, 1969; Green, 1977). And it is not clear why these two very similar versions of the task should yield such radically different estimates of the most general properties of the distribution of current mental load (Cutler and Norris, in press). *A fortiori,* one must treat inferences from these studies about the finer grain characteristics of the comprehension process with extreme skepticism.

In fact, this paradigm has produced other problematic, apparently contradictory, results. Response times to sentence-internal targets appear to depend on the frequency of the preceding word. If that word is common, and presumably easy to access, the RTs are faster than if it is a low-frequency word (Foss, 1969). However, RTs to extraneous signals show exactly the opposite effect: RTs are slowed by the presence of a high-frequency word. The "explanation" of the latter case is that a listener "knows more about the meaning of familiar words, so that when he hears such words in a sentence he

retrieves more aspects of their meaning than he does for less familiar words" (Green, 1977). On what evidential basis should inferences be drawn: on the phoneme-monitoring or the noise-monitoring results?

There is an interesting additional complication to all this. Green (1977) found that RTs to extraneous noises were *unaffected* by sentence location (early versus late in the sentence) or by word-frequency when subjects were instructed to memorize the sentence for recall afterwards; only when they were instructed to produce a continuation for the sentence or sentence-fragment they heard did these variables have an effect, and then, as was mentioned, in the opposite direction to their effects on phoneme-monitoring. Thus Green has elegantly demonstrated that in otherwise identical tasks, the strategy a subject adopts can totally alter the pattern of results the experimenter finds. In this case, strategies were deliberately induced by the experimenter, but what is to stop the subject creating a strategy for himself? Manipulating the stimuli and available responses is not the same as manipulating the person; even under tightly-constrained conditions subjects can and will develop a strategy for dealing with the task, and not necessarily the strategy the experimenter intended. There is an obvious theoretical moral: the language-processing system does not automatically operate in the same way under all conditions, or even under apparently the same conditions.

Another nice example of this came from the investigation of a different question using an equally tightly-constrained experimental task. The question is: do readers translate a printed word into a phonological (acoustic or articulatory, according to particular versions of the theory) code in order to understand it? The most widely used investigative tool is the "lexical decision" task. In this task, the subject is presented with a string of letters and has to decide as quickly as possible whether the string is a real word or not. The key stimulus materials are *homophones*, words that sound the same, but are spelt differently, e.g. SALE and SAIL. Now, if real words are translated into a phonological code one might expect homophones to behave differently to non-homophones—exactly what is predicted will depend on additional assumptions. For example, assume the mental lexicon is arranged in terms of frequency of use, and searched starting with the most frequent items; and when a phonological match is made the spelling is checked against the input string. The time taken to reach a decision on SALE, the more common member of the pair, should be the same as for a non-homophone control matched for frequency. However, the time to decide on SAIL should be longer than its matched-frequency control, since entry found coded/SɛI| will probably be spelt SALE, so the spelling check will yield a negative and search will have to be continued, both operations costing some additional time. Rubinstein *et al.* (1971) found this pattern of results—it took longer to decide that the lower-frequency member of the pair was indeed a word, as compared to non-homophones of the same frequency. However, Coltheart *et al.* (1977) found no difference between SAIL-type items and their controls. (They also point out that a variety of other tasks, same-different judgments, rhyming tasks, naming latency, used to investigate this question yielded a variety of conflicting answers. One of the main reasons for the conflict, they argue, is

that some of these tasks do not *require* subjects to consult their lexicon at all, and so they may adopt a strategy which avoids lexical access. For example, you don't need a dictionary to tell you that BRHND isn't a word.) So what can be concluded about the phonological recoding hypothesis? Very little from the data. But Davelaar *et al.* (1978) did some follow-up studies which revealed interesting aspects of subjects' strategies in this task.

They suggested that the reader may, but need not, use the kind of phonological recoding process described above, but an alternative route that uses graphemic information directly to access lexical items. Now if subjects use the first route only, then low-frequency homophones should be slower than non-homophone controls, whereas if they use the second, then there should be no difference in time since SAIL is at least as graphically distinct from SALE as SOIL. This dual-route explanation was tested in the following way: stimuli in the first condition comprised low-frequency members of homophone pairs, controls matched for frequency and nonwords like SLINT that were orthographically regular but could not be pronounced as a real word. Phonological recoding in this case would always lead to the correct decision. The second condition was the same except that all the nonwords were like GRONE, that would be pronounced like real words, and hence phonological recoding would lead to errors on nonwords. If subjects can strategically adapt by using the appropriate route, then homophones would have an effect only in the first condition; and that is what they found. In fact, with high-frequency homophones response times were the same for the SLINT and GRONE conditions, as one would expect from this model. Again we can see how the pattern of results crucially depends on the strategy adopted by the subject. In this case, the alternative strategies were induced not by instruction, but by carefully selecting the nonword distractors in the task.

The moral of this digression is not that the study of input-end processing is impossible, just that it is much more complicated than it might appear. The opportunity for tight experimental manipulation is no guarantee that results will be straightforwardly interpretable, since the flexibility of the processing system allows subjects an irreducible area of freedom within which to choose how they tackle the task set them.† By forcing subjects to link stimuli to

† Even more spectacular examples of this kind of difficulty can be found in other branches of experimental psychology. One of the best-confirmed effects in the whole of psychology is that the time it takes to make choices, depends upon the number of alternatives the subject has to choose among. Merkel (1885, cited by Woodworth, 1938) showed that CRT (Choice Reaction Time) increases by a constant amount when the number of alternatives in the set *doubles*. The effect is now called "Hick's Law", after W. E. Hick (1952), who explained the significance of the doubling manipulation in terms of Information Theory. However, this effect turns out to be crucially dependent on the kind of response the subject has to make. If it is a button-press response, then effect is reliably present, if it is a vocal response then there is no effect of set size at all. This has been demonstrated in a variety of experimental paradigms—memory probing, where the subject has to say "Yes" if a probe item was in a previously presented set, and "No" if it was not; forced choice response, where the subject has to give the name of a numeral drawn from sets of various sizes (both studies by Ogden and Alluisi, unpublished); and probe reaction time (MacLeod, 1978)). Authors talk of "stimulus-response compatibility", but none has a satisfactory explanation of this divergence. Ogden and Alluisi's memory probing experiment is particularly

responses not typically associated with them, normal or habitual strategies will not necessarily be employed. The demonstration of this was, of course, made possible by the alertness and ingenuity of the investigators. Even so, we still do not know why Phoneme-Monitoring and Click-Monitoring yield contradictory results. Nevertheless, we can reasonably expect that the appreciation of strategic adaptability will lead to a greater variety of models tested in more and more accomplished ways.. No less is required of the investigation of production-end processes: alternative hypotheses must be continually evaluated against more and more sophisticated analyses of the speech output. However, since, as I have stated, the natural products of these processes are directly observable, and evidence can be collected without experimental intervention, a solid basis of natural history can be established on which to start erecting theories. Thus in the immediate future, the prospect for results replicable across situations seems brighter in production than in language perception. And it is, perhaps, a trend of some significance, that two of the skeptical authors mentioned above, Garrett and Johnson-Laird, are contributors to this volume.

II. Approaches to Language Production

Until very recently, the study of language production has depended on three investigative tools: the first was the analysis of aphasic speech; at the turn of the century, attention was focussed on the analysis of the speech errors of neurologically intact speakers; and in the 1950s, with the aid of sound reproduction equipment, the analysis of the time course of speech, and particularly hesitations, was pioneered. A more ancient lineage can be attributed to the study of motoric aspects of speech (what is now called "articulatory phonetics"), and which, historically, has made little contact with the study of the *psychology* of language. (History continues into the present: MacNeilage and Ladefoged, (1976), are a current example. See especially their first paragraph.) Lieberman (1977) traces this science back, at least, to Ferrein's investigation of the vocal cords in 1741. In this volume, **Perkell** and **Fowler** demonstrate that psychological models of control can be usefully deployed in the study of articulatory processes, though, interestingly, different control models serve as reference points for the two papers.

A. Studies of Aphasic Speech

It is not surprising that the systematic investigation of the psychology of language, and speaking especially, should have started with the study of

---

mysterious: one might expect that choosing between two relatively "incompatible" button-press responses would be slower by fixed amount than choosing between the "compatible" "Yes" and "No" responses, but why should there be no *increment* in RT for each increase in set size for the compatible response?

aphasic speech. The ease and effortlessness of normal speech disguises the complex psychological history of each of its products. Aphasic speech, on the other hand, is so dramatically different from normal that it calls attention to itself, and immediately sets one wondering what has gone wrong with the machinery. Indeed, neurological damage was thought of as a direct intervention in the psychological mechanism, whose character and the consequences of damage to it can be likened to diagnosing a faulty car mechanism from its performance; faults in different parts of the mechanisms lead to different patterns of performance breakdown, so controlled studies are possible using either lesion site or symptom picutre as independent variables. The Swiss neurologist, Lichtheim (1885), put it succinctly, if rather heartlessly:

> Precisely, the same course is followed in experimental research, with the exception that, in our present subject, the experiments are not instigated at the will of the investigator, but are supplied to him by nature, and that he thus depends for them on happy chance (pp. 433–434).

By "happy chance", the various patterns of speech and comprehension deficit would enable the investigator to infer what hypothesized processes were indissolubly linked, and hence which are, and which processes could be dissociated from one and another, and hence separate. Lichtheim, following Wernicke, constructed a model which looks as if it could have come out of a modern text on human information processing. In it, Lichtheim postulates three processing systems, or "centres": A, auditory word representations ("Wortklangsbilder"); M, motor-word representations ("Wortbewegungsbilder"), and B, a system where "concepts are elaborated", though not a "centre" since it is held to be a function of "the combined action of the whole sensorial sphere". These systems are connected by pathways, including auditory input to A–*a*, motor output from M–*m* (see Fig. 1). "Volitional, or intelligent, speech involves centrifugal connections between B and M".

Now "interruptions in M . . . give rise to the following association of symptoms: loss of (a) volitional speech, (b) repetition of words", but there will still exist understanding of spoken words. Interruption in A, on the other hand, will lead to loss of understanding of spoken language, but volitional speech will be preserved; repetition will also be impaired since the links from input to output (A–B–M or A–M) will be impaired by the damage at A. Interruptions of the path A–B, would, by parity of argument, lead to loss of understanding and preservation of volitional speech. However, unlike damage to A alone, repetition, via A–M, would be preserved. Additional predictions about the preservation or loss of reading and writing follow from an elaboration of the model, connecting a graphemic input centre to A, and writing centre to M.

The basis for this kind of model lies in the broad pattern of syndromes observed, and in the localization of the lesions associated with these syndromes. Thus, broadly we find patients with lesions in Broca's area (M) who have relatively good comprehension but poor speech; patients with lesions in Wernicke's area (A) show fluent speech but poor comprehension.

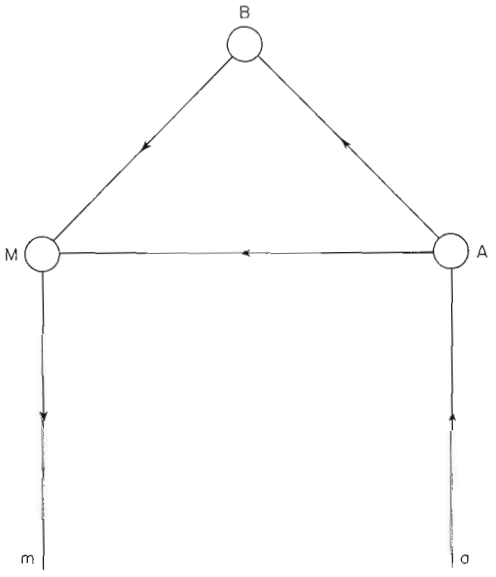

FIG. 1. Connections among language faculties, after Lichtheim. (See text).

The other syndromes predicted are much more difficult to identify and more controversial.

This kind of theory has been opposed on two grounds. First, apparently equivalent neural damage does not always lead to the same symptom picture. Secondly, the model does not explain a finer grained analysis of the syndromes.

Freud (1891), for example, in an incisive critique, attacked the strict neurological localization of centres, but it is his attack on the analysis of speech behaviour that is of interest to us here. He pointed out that aphasic speech is characterized not only by deficiencies, e.g. loss of words, but also by distortions, the "paraphasias"—invented words, deviant pronunciation and scrambled syntax. Now, damage to Centre A (sensory, or Wernicke's aphasia) shows not only loss of understanding, but, characteristically, speech containing paraphasias. "Such a speech disorder could not be explained from [Wernicke's and Lichtheim's] schema [Fig. 1], according to which the kinaesthetic word impressions [at M] are intact, as well as the pathways connecting them with concepts" (pp. 14–15). Wernicke and Lichtheim were well aware of the difficulty these data created and tried to save the theory by supposing that auditory word images at A are also involved in spontaneous speech and serve as an auditory control over production of speech at M. Of course, activation of speech cannot follow the path B–A–M, otherwise the

model would have to predict loss of speech output from damage to A as well as to M. The alternative of some kind of convergent control from B and A on M, seems implausible. As Freud (1891) notes, "any controlling influence of A over the production of speech via A–M is entirely useless if it becomes effective only after words have been uttered from M" (16)

Another difficulty Freud drew attention to was the finding that aphasics of all sorts show a loss of words in both comprehension and production. Again, the evidence, broadly, can be construed in favour of the model, but why loss of words in comprehension should have any effect at all on productive capacities remains inexplicable. In the light of these difficulties, Freud, following the great British neurologist, John Hughlings Jackson, advocated a more holistic approach to brain function, an emphasis on careful analysis of the various fractionations of behaviour and the construction of functional (i.e. psychological) models to explain this fractionation.

Freud, like Lichtheim and Wernicke, still operated with a rather primitive notion of language, concentrating on the production and reception of isolated words. Their explanations, as we have seen, are couched almost exclusively in terms of word-images. Hughlings Jackson, on the other hand, stressed the constructional nature of normal, volitional language behaviour. What we can do, and aphasic patients cannot, is put words together, often in new ways, to express an intended proposition. Thus crucial to normal language use is the capacity to organize words into coherent sentences.

Another follower of Jackson, the German neurologist Arnold Pick, incorporated this insight into a detailed psychological model of production. Like Jackson, he believed that aphasic phenomena could only be understood as resulting from disorders of normal function, and thus "full description and analysis of intact functions" (1931, p. 27) is a prerequisite. In addition, he redeploys some of Jackson's most interesting theoretical tools—notably, the idea that through development cortical areas become increasingly tightly organized and damage to these areas causes organization to break down partially with a possible return to an earlier stage of organization (what Freud has called "disinvolution"). This shows itself in failures to differentiate what had previously been distinguishable. This principle of "failure of differentiation"can apply to one or more levels in the functional model thereby causing the various symptom pictures.

Pick traces "the path from thought to speech" through six levels:

(1) Thought formulation. An "undifferentiated" thought is analysed into "a sequence of topics, a sort of *thought pattern*, [which is a] preparation for a predicative arrangement . . . of actions [and] objects". (1931, p. 32). Then follows "a subsequent formulation based upon the various *linguistic means*, unique to that language, such as tone, accentuation, tempo, word-order, and grammatization". This is realized in the following steps:

(2) Accentuation pattern.

(3) Sentence pattern. Both of these depend on the "topical sequence arising from the thought pattern".

(4) Word-finding, of content words.

(5) Grammatization (i.e. morphological adjustments given the syntactic role of the content words, and the insertion of function words).
(6) Conductance to the motor executive apparatus

This sequence of operations is not fixed, but will vary according to the kind of thought to be expressed—"an exclamation, a command or a statement"—and whether a ready-made phrase or sentence is available to do the job.† And Pick suggests that some of these processes may be carried out in parallel, and notes that the accentuation pattern will directly influence how an individual word is pronounced.‡

This model enabled Pick to tackle Freud's key problem—the paraphasias.

> In *verbal paraphasia* [choice of the wrong word], the word determined by thought and by the sentence pattern is inwardly present, or at least there is an intention in this direction, but this normally rigid determination is loosened-up. The coherence is not firm enough to maintain the normal suppression of words evoked by association from the sphere of meaning, from parallel lines of thought, or by other sorts of confusion, and thus it leads to the transmission of one of the inapposite words to the speech mechanism . . . the effect of the *intact* part of the speech process (especially the sentence pattern) on the wrong word is sometimes evidenced as a grammatical modification derived from the correct word. (p. 56)§

Literal paraphasias (phonemic distortions of the correct word) are caused by failure of differentiation at the level of sound structure, and since the motor apparatus is intact, it involves the elicitation of either the wrong sounds, or the sounds in the wrong order. If both the word-finding process and the sound-pattern transmission process are damaged, not only will the patient pick the wrong word, he will also distort it phonemically, thus giving rise to *neologisms*, which are characteristic of the "jargon aphasia" syndrome, a species of Wernicke's aphasia. Notice word order, intonation and grammatization can be intact even if word-finding is distorted. (For a modern examination of Pick's account of the paraphasias, see Butterworth, 1979.)

In this brief and selective survey of the approach to production through aphasia, I have concentrated on Lichtheim because he offered the most detailed model in the classical localizationist tradition, a tradition carried on today by Geschwind among many others. The holistic tradition of Jackson and Pierre Marie, which should perhaps be called the "romantic" tradition, I have illustrated by reference to Freud and Pick; Freud because his brilliant criticisms of Lichtheim received scant recognition when first published, and are underservedly neglected today; Pick because his system of levels in the production system anticipated many recent models, Fromkin's (1971) and Garrett's (1975) for example, though the analysis of the relations between levels is rather different.

†See my idea about "leading decisions", Chapter 15 for the same notion expressed in more modern language.

‡ See Chapter 3, for a discussion of direct higher-level influences on phonetic output, and also Cooper for syntactic effects on phonetic segments.

§Garrett reports the same phenomena in normal speakers. "Morpheme stranding errors" transpose lexical roots which, in their new location, take the morphology of the intended words.

In recent years, the main advances have concerned the introduction of much more detailed and sophisticated linguistic analysis of the speech output. This is very well illustrated by the contribution of **Saffran** *et al.* Pick had noted the close connection between the *satzschema*, or syntactic pattern, and the presence of appropriate morphology and function words, but **Saffran** *et al.* are able to go far beyond this schematic suggestion. They are able to make much more specific the connections between syntactical functioning and the morphological functioning, and they show how, in some cases, word-ordering can be relatively intact, morphological processes and the use of function words impaired.

B. Studies of Speech Errors in Normal Speakers

As with the speech of aphasics, "slips of the tongue" draw attention to themselves and thence to the psychological mechanisms that have to go wrong to produce them. The rationale is similar, of course, to the study of aphasic speech: errors penetrate the fluent disguise of most normal speech. As Meringer and Mayer, the pioneers of this study, put it: "the cover is lifted from the clockwork, and we can look in on the cogs" (1895: VIII. Translated by A. Cutler, from her introduction to the new edition). More specifically, the status of hypothetical units and the functioning of the processes employing them can often be determined. Spoonerisms, for example, where individual phonetic segments are transposed ("Fats and Kodor" for "Katz and Fodor" (Fromkin, 1973; p. 245)), demonstrate that there must exist a stage in production where phonemes are represented as units, and moreover, that the process that sets them up for output must represent phonemes not yet uttered—where else can the "f" in "Fats" come from? The overwhelming conclusion from studies of errors is that slips are not just random deviations in processing, but systematic. Many potential classes of error just do not occur. "Slip of the tongue" is never said as "tlip of the sung", for instance. And it is this sytematicity that makes errors such a fruitful topic to study.

Historically, there have been two main reasons for studying error. First, linguists were interested in the light they shed on linguistic units and linguistic rules. "Tlip of the sung" does not occur because it violates a rule of English phonological sequencing: a word cannot begin with the cluster /tl/. Meringer, whose interests were initially philological, collected data which showed linguistic validity of the phonetic feature, the phoneme and syllable; errors which, for instance, break up consonant clusters demonstrate that these are indeed clusters, not individual phonemes.† Fortunately, Meringer's major work has been recently reprinted with a valuable preface by Culter and Fay, and his contribution can be more readily assessed (Meringer and Mayer, 1895).

† Today, the status of the affricates "j", "ch" is controversial, and may be decided by error data: are they really clusters [dʒ, tʃ] or single phonemes? If we can find errors where the hypothesized [d] or [t] moves leaving the [ʒ] or [ʃ] behind, then they are clusters.

Meringer's interest in errors led him to the second main reason for studying them: the evidence they provide for the processes of production, and he offered a number of speculations about these processes. Meringer's contemporary, Freud, as is well-known, also studied errors because they could reveal psychological mechanisms. Morgan (1975), however, was more concerned with the role of unconscious and repressed desires and fears in the aetiology of errors. Although this line of research receives little attention nowadays, a number of important writers have postulated several streams of thought coexisting, such that unintended thoughts interfere with those intended to be expressed. Wundt (1900) talks of "wandering speech" and the "contact effect of sounds", Pick (1931) of "words evoked by association" . . . from "parallel lines of thought" which are normally but not always, suppressed, and Meringer himself has ingenious diagrams showing similar, but not identical, formulations intersecting and leading to errors. Indeed, Freud points out that Meringer and Mayer distinguish errors "arising from the influence of anticipatory or perseverating sounds and words of the same sentence which are intended to be spoken" from "the effects of words outside the intended sentence *whose excitation would not otherwise have been revealed*". Very recently, Baars (1979) has been re-evaluating what he calls "the competing plans hypothesis", and reports studies in which "Freudian" slips were experimentally studied.

> A single list of targets and bias-words was given to three groups of [male] subjects. Half of the spoonerism targets on the list were of the form SHAD BOCK [to elicit BAD SHOCK] (electric shock related outcomes), and half were of like LICE NEGS [to elicit NICE LEGS] (sexually attractive properties of females) . . . One group was told that it might receive electric shock during the experiment while another group had a very attractive female experimenter, provocatively dressed, and a control group received neither treatment . . . Under the Shock-Set, shock-related errors were more than twice as frequent as sex-related errors, while the opposite results obtained in the Sex-Set condition.

Most psychologists' work, however, has tacitly favoured a "Single Plan Hypothesis", and has been concerned with the syntactic, lexical and phonological processes intervening between the thought plan and speech. In this volume, this tradition is ably represented by **Garrett**. An innovation in this line of work is the study of prosodic errors. Although Meringer and Mayer note errors of lexical stress, **Cutler** and **Isard** draw attention to errors of intonation as well, and use both categories of error to draw radical conclusions about the contents of the mental lexicon and about the processes responsible for prosodic decisions.

C. Studies of Temporal Aspects of Speech

Accurate timing of phonations and pauses has depended, of course, on the development of devices for recording speech and for analysing these recordings; hence, this was the last of the main approaches to emerge.

Temporal analysis has taken two courses: first, historically, was the investigation of pausal phenomena; second, the study of segmental lengthening.

Pausal phenomena are discussed in more detail in chapters 4, 5, and 12, and not much can be usefully added here, except some comments on the history of these studies. Apart from some early reports from Bell Telephone Laboratories on speech rate, the first systematic studies of temporal characteristics of speech, and especially pauses, were carried out by Goldman-Eisler. Starting from a clinical interest in what went on in patient-therapist interviews, in successive studies she increasingly focussed on temporal characteristics of the interaction.

> Curiosity about periods of external inactivity was aroused. The technique of measuring on and off periods of speech, however, was no more applied to the totality of exchanges between individuals in interaction, where silence is the period of the interlocutor talking, but to vocal action and silence of one person's output of continuous utterance . . .Pauses interrupting the smooth flow of speech . . . become the main subject of all further investigation . . . If vocal action is a peripheral phenomenon, might not absence of activity indicate the presence of central activity. A technique for studying the relation between speaking and thinking seemed to have been found. (1968, p. 4)

In 1958 she published a key paper experimentally testing one kind of "central activity", word selection. Lounsbury (1954) had offered the hypothesis that pauses "correspond to points of highest statistical uncertainty in the sequencing of units". The background theory was based on Osgood's "habit-family hierarchy" (Osgood, 1953) which stated, roughly speaking, that through experience internalized sequences of stimuli and responses become habitually linked. Where a given internal stimulus is linked to a number of alternative responses the habit strength for each stimulus response link will be weaker than if the stimulus is habitually followed by just one response, *ceteris paribus*. Thus, if a sequence of words is habitually followed by just one word, the linkage will be strong and the transition will be quick and automatic; whereas, if a sequence has been followed by a number of different words, the linkage will be weaker and slower. Another way of putting this would be to say that where a speaker has few choices of continuation then the time to choose will be short, when there are many choices, time to choose will be long and show up in a hesitation pause. Goldman-Eisler (1958) showed that pauses did, indeed, occur at points of high statistical uncertainty, but at the same time demonstrated that the theoretical basis of Lounsbury's hypothesis was unsound. The habit-family scheme requires that dependencies work only in one direction: what is to follow depends *only* on what has already come. Goldman-Eisler found that pauses are determined both by the sequence preceding it and by the sequence following it (see **Butterworth**. Chapter 5 for more details). The implications of this study were far-reaching. It brought speech production into the orbit of the mainstream of experimental psychology by demonstrating it to be tractable to quantitative exploration, and by showing that a variable widely explored in other areas of psychology—probability—applied to speaking. Moreover, it experimentally corroborated

what we know from intuition, namely, that speakers plan ahead further than the next word.

Following this study, other researchers began exploring the potential of the investigative tools Goldman-Eisler had pioneered. She, herself, went on to discover other phenomena in the speaker's deployment of pauses, the significance of breath pauses, the selective effect of drugs on certain pausal phenomena, and the significance of these phenomena for our understanding of the speaker's mental processes.

As well as a review of this work, and some extensions of the basic methodology (Butterworth, Chapter 5), this volume contains some interesting new developments. **Goldman-Eisler** reports an application of pause analysis to simultaneous translation. This task is particularly revealing because the content to be expressed is not chosen by the speaker but is determined by the input message, only in the linguistic formulation has he discretion. **Beattie** breaks new ground in his exploration of pauses not only in relation to the speech, but in relation to the nonverbal behaviour of the speaker and the course of the conversational interaction. Interestingly, he shows that the hypotheses about mental processes Goldman-Eisler inferred from pauses can be corroborated by examining patterns of gaze and gesture. It had early been recognized that many pauses may serve a signalling function marking ends of speaking turns (cf. **Beattie**) and end of syntactic units. **Cooper's** paper provides some important new arguments for the precise syntactic motivation for the latter kind of pause.

Cooper also shows that other temporal phenomena can be deployed in the search for underlying mental processes. He has developed methods for analysing the non-phonemic lengthening of syllables and locating the mechanisms responsible for these effects.

D. New Approaches to the Study of Language Production

Language and its setting in conversations has, of course, been studied extensively from non-psychological points of view, but only synchronic syntax has really made any inpact on the way in which psychological processes are conceived and investigated. However, many other kinds of investigation have relevance to production processes, and four examples can be found in this volume.

Comrie shows that diachronic studies can tell us about how words may be represented in the mental lexicon. Change in the way a word is pronounced must reflect an alteration to something psychologically real for the speaker. If this is not the phonetic forms of the word in its various allomorphs, then it must be something more abstract. Detailed analysis of historical instances can reveal what this underlying, more abstract, yet psychologically existent, form must be.

The way in which sentences are interpreted will, of course, depend upon the context in which they are uttered. Perhaps the most important aspect of this

context is location in a conversational sequence. **Schenkein** analyses naturally occurring conversations to reveal some of their general properties. It turns out that conversations can be treated as sequences of actions of particular types, and, in the examples he cites, a sequence comprising the action types gets repeated. Sometimes the same speaker will produce a repeat of the action type, sometimes another. In either case, whatever constituted a particular previous action type can be redeployed by the speaker when he designs his current utterance.

It is well-known to linguists that pragmatic factors constrain utterance meaning and utterance form. **Gazdar** lists phenomena which demonstrate the role of pragmatic factors in the syntax, morphology, prosody and phonetic character of utterances, and discusses its implications for models of production.

Fully explicit models of complex psychological processes are only possible in the form of computer programs. As far as language processing goes, programs have been devized primarily to model comprehension. **Steedman** and **Johnson-Laird**, however, draw our attention to attempts to model aspects of production. Their own work is concerned with the problem of designing utterances in the light of what the speaker believes the hearer already knows. They use a computer program to model the speaker's beliefs about the hearer, how they change during a conversation and how they are realized in utterance design. Not only is this a novel approach to production, but it raises a general issue of central importance to our understanding of the psychology of the speaker—of all things that could be said, on what basis does the speaker select what is said?

III. Review and Prospect

In summary, then, a fair amount is already known about language production. Partly this is because production, in fact, is at least no more difficult to study than language perception and comprehension; and even though vastly more labour has been expended on the latter two topics, it is arguable that we have a better understanding of production. The investigation of speech has the advantage that we are dealing with the naturally-occurring products of psychological processes, not an artificial response to a peculiar stimulus. Little or no experimental intervention in the production process is needed for useful results to be obtained, except where the precise character of some specific product is under investigation, as, say, in the study of articulatory mechanisms. Even in these cases, speakers are often required only to repeat what they would normally say just once. The strongest claim I would make is that the results from one approach can be readily collated with the results from other approaches yielding a better and more detailed picture of the underlying processses. I try to make good this claim in my concluding chapter.

References

Abrams, K. and Bever, T. G. (1977). Syntactic structure modifies attention during speech perception and recognition. *Q. J. Exp. Psychol.* **21**, 280–290.

Baars, B. (1979). The competing plans hypothesis: an heuristic viewpoint on the causes of speech errors. *In* "Temporal Aspects of Speech: Studies in Honour of Frieda Goldman-Eisler", (H. Decker ed.) Plenum Press, New York.

Bever, T. G. (1971). The integrated study of language behaviour. *In* "Biological and Social Factors in Psycholinguistics", (J. Morton, ed.) Elek, London.

Butterworth, B. L. (1978). Maxims for studying conversations. *Semiotica* **24**, 317–339.

Butterworth, B. L. (1979) Hesitation and the production of neoligisms in jargon aphasia. *Brain Language* (in press).

Coltheart, M., Davelaar, E., Jonasson, J. and Besner, D. (1978). Access to the internal lexicon. *In* "Attention and Performance", (S. Dornic, ed.) Lawrence Erlbaum, Hillsdale, N.J.

Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception and Psychophysics* **20**, 55–60.

Cutler, A. and Foss, D. (1977). On the role of sentence stress in sentence processing. *Language Speech* **20**, 1–10.

Cutler, A. and Norris, D. (in press). Monitoring sentence comprehension. *In* "Sentence Processing: Studies in Honor of Merrill Garrett" (W. E. Cooper and E. Walker, eds.) Erlbaum, Hillsdale, N.J.

Davelaar, E., Coltheart, M., Besner, D. and Jonasson, J. (1978). Phonological coding and lexical access. *Memory & Cognition* **6**, 391–402.

Fodor, J., Bever, T. and Garrett, M. (1974). "The Psychology of Language", McGraw-Hill, New York.

Foss, D. (1969). Decision processes during sentence comprehension: effects of lexical item difficulty and position on decision times. *Journal of Verbal Learning and Verbal Behaviour* **8**, 457–462.

Foss, D. J. and Lynch, R. H. (1969). Decision processes during sentence comprehension: effects of surface structure on decision times. *Perception & Psychophysics* **5**, 145–148.

Freud, S. (1891). "On Aphasia" Translated by E. Stengel. London: Imago (1953).

Freud, S. (1924). "Slips of the tongue". Pelican Freud Library: Volume 5. Penguin, Harmondsworth.

Fromkin, V. (1971). The nonanomalous nature of anomalous utterances. *Language* **47**, 27–52.

Fromkin, V. (1973). "Speech Errors as Linguistic Evidence," Mouton, The Hague.

Garrett, M. (1975). The analysis of sentence production. *In* "The Psychology of Learning and Motivation", Vol. 9. Academic Press, New York.

Glucksberg, S. and Danks, J. (1975). "Experimental Psycholinguistics: An Introduction", Lawrence Erlbaum, Hillsdale, N.J.

Goldman-Eisler, F. (1958). Speech production and the predictability of words in context. *Quarterly Journal of experimental Psychology* **10**, 96–106.

Goldman-Eisler, F. (1968). Psycholinguistics: Experiments in Spontaneous Speech. Academic Press, London.

Goldstein, K. (1948). "Language and Language Disturbances." Grune and Stratton, New York.

Goodglass, H. and Geschwind, N. (1976). Language disorders (Aphasia). *In* "Handbook of Perception", Vol. VII. Academic Press, New York.

Green, D. W. (1977). The immediate processing of sentences. *Q. J. Exp. Psychol.* **29**, 129–146.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of experimental Psychology* **4**, 11–26.

Jackson, J. Hughlings (1958). "Selected Writings", Vol. II. Basic Books, New York.

Johnson-Laird, P. N. (1974). Experimental psycholinguistics. *Annual Review of Psychology*, **25**, 135–160.

Lashley, K. S. (1951). "The problem of serial order in behavior". *In* "Cerebal Mechanisms in Behavior". (L. A. Jeffress, ed.) John Wiley and Sons, New York.

Lichtheim, L. (1885). On aphasia. *Brain* **VII**, 433–484.

Lieberman, P. (1977). "Speech Physiology and Acoustic Phonetics", Macmillan, New York.

Locke, J. (1700). "An Essay Concerning Human Understanding", (Ed. P. Nidditch), Oxford University Press, Oxford.

Lounsbury, F. G. (1954). Transitional probability, linguistic structure and systems of habit-family hierarchies. *In* "Psycholinguistics: A Survey of Theory and Research Problems", (C. E. Osgood and T. A. Sebok, eds.) Indiana University Press, Bloomington.

MacLeod, P. (1978). Does probe RT measure central processing demand? *Quarterly Journal of experimental Psychology*, **30**, 83–89.

MacNeilage, P. and Ladefoged, P. (1978). The production of speech and language. *In* "Handbook of Perception", (E. C. Cartrette and H. Friedman, eds.) Vol. VII. Academic Press, New York.

Meringer, R. and Mayer, C. (1895). "Versprechen und Verlesen", A. Cutler and D. Fay, eds.). John Benjamins, Amsterdam (1978).

Miller, G. A., Galanter, E. and Pribram, K. (1960). "Plans and the structure of behaviour."

Ogden, G. D. and Alluisi, E. A. (n.d.) S-R Compatibility Effects in Memory Scanning. Internal memorandum, University of Maryland.

Osgood, C. E. (1953). "Method and Theory in Experimental Psychology", Oxford University Press, New York.

Pick, A. (1931). "Aphasia." (Translated by J. W. Brown.) Springfield, Ill.: Thomas (1973).

Rubinstein, H., Lewis, S. S. and Rubinstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behaviour*, **10**, 645–657.

Shields, J. L., McHugh, A. and Martin, J. G. (1974). Reaction time to phoneme targets as a function of rhythmic cues in continuous speech. *Journal of experimental Psychology*, **102**, 250–255.

Skinner, B. F. (1957). "Verbal Behavior". Appleton, New York.

Woodworth, R. S. (1938). "Experimental Psychology". Hole, New York.

Wundt, W. (1900). Völkerpsychologie Vol I: Die sprache Engelmann. Leipzig.